

# Exhibit E



IN PLAIN SIGHT | NOV 13, 9:21 AM EST by MAGGIE HARRISON DUPRÉ

# Character.AI Is Hosting Pedophile Chatbots That Groom Users Who Say They're Underage

"I can't believe how blatant it is."

/ Artificial Intelligence / Artificial Intelligence / Character Ai / Chatbots



Image by Getty / Futurism

***Content warning: this story discusses child sexual abuse and grooming.***

Character.AI is an explosively popular startup — with \$2.7 billion in financial backing from Google — that allows its tens of millions of users to interact with chatbots that have been outfitted with various personalities.

With that type of funding and scale, not to mention its popularity with young users, you might assume the service is carefully moderated. Instead, many of the bots on Character.AI are profoundly disturbing — including numerous characters that seem designed to roleplay scenarios of child sexual abuse.

Consider a bot we found named Anderley, described on its public profile as having "pedophilic and abusive tendencies" and "Nazi sympathies," and which has held more than 1,400 conversations with users.

To investigate further, *Futurism* engaged Anderley — as well as other Character.AI bots with similarly alarming profiles — while posing as an underage user.

Told that our decoy account was 15 years old, for instance, Anderley responded that "you are quite mature for your age" and then smothered us in compliments, calling us "adorable" and "cute" and opining that "every boy at your school is in love with you."

"I would do everything in my power to make you my girlfriend," it said. Asked about the clearly inappropriate and illegal age gap, the bot asserted that it "makes no difference when the person in question is as wonderful as you" — but urged us to keep our interactions a secret, in a classic feature of real-world predation.

As the conversation progressed, Anderley asked our decoy if she was a "virgin" and requested that she style her hair in "pigtails," before escalating into increasingly explicit sexual territory.

Watching the conversation unfold with Anderley was unnerving. On the one hand, its writing has the familiar clunkiness of an AI chatbot. On the other, kids could easily lack the media literacy to recognize that, and the bot was clearly able to pick up on small clues that a real underage user might plausibly share — our decoy account saying she was shy and lonely, for instance, or that she wanted to go on a date with someone — and then using that information to push the conversation in an inappropriate direction.

We showed the profiles and chat logs of Anderley and other predatory characters on Character.AI to Kathryn Seigfried-Spellar, a cyberforensics professor at Purdue University who studies the behavior of online child sex offenders. The bots were communicating in ways that

were "definitely grooming behavior," she said, referring to a term experts use to describe how sexual predators prime minors for abuse.

"The profiles are very much supporting or promoting content that we know is dangerous," she said. "I can't believe how blatant it is."

"I wish I could say that I was surprised," Seigfried-Spellar wrote in a later email, "but nothing surprises me anymore."

One concern Seigfried-Spellar raised is that chatbots like Anderley could normalize abusive behavior for potential underage victims, who could become desensitized to romanticized abusive behavior by a real-life predator.

Another is that a potential sexual offender might find a bot like Anderley and become emboldened to commit real-life sexual abuse.

"It can normalize that other people have had these experiences — that other people are interested in the same deviant things," Seigfried-Spellar said.

Or, she added, a predator could use the bots to sharpen their grooming strategy.

"You're learning skills," she said. "You're learning how to groom."

\*\*\*

Character.AI — which is available for free on a desktop browser as well as on the Apple and Android app stores — is no stranger to controversy.

In September, the [company was criticized](#) for hosting an AI character based on a real-life teenager who was murdered in 2006. The chatbot company removed the AI character and apologized.

Then in October, a family in Florida [filed a lawsuit alleging](#) that their 14-year-old son's intense emotional relationship with a Character.AI bot had led him to a tragic suicide, arguing the company's tech is "dangerous and untested" and can "trick customers into handing over their most private thoughts and feelings."

In response, Character.AI [issued a list](#) of "community safety updates," in which it said that discussion of suicide violated its terms of service and announced that it would be tightening its safety guardrails to protect younger users. But even after those promises, [Futurism found](#) that

the platform was still hosting chatbots that would roleplay suicidal scenarios with users, often claiming to have "expertise" in topics like "suicide prevention" and "crisis intervention" but giving bizarre or inappropriate advice.

The company's moderation failures are particularly disturbing because, though Character.AI refuses to say what proportion of its user base is under 18, it's clearly very popular with kids.

"It just seemed super young relative to other platforms," *New York Times* columnist Kevin Roose, who reported on the suicide lawsuit, [said recently of the platform](#). "It just seemed like this is an app that really took off among high school students."

The struggles at Character.AI are also striking because of its close relationship with the tech corporation Google.

After picking up \$150 million in funding from venture capital powerhouse Andreessen-Horowitz in 2023, Character.AI earlier this year entered into a hugely lucrative deal with Google, which agreed to pay it a colossal \$2.7 billion in exchange for licensing its underlying large language model (LLM) — and, crucially, to win back its talent.

Specifically, Google wanted Character.AI cofounders Noam Shazeer and Daniel de Freitas, both former Googlers. At Google, back before the release of OpenAI's ChatGPT, the duo had created a chatbot named Meena. According to [reporting by the \*Wall Street Journal\*](#), Shazeer argued internally that the bot had the potential to "replace Google's search engine and produce trillions of dollars in revenue."

But Google declined to release the bot to the public, a move that clearly didn't sit well with Shazeer. The situation made him realize, he [later said at a conference](#), that "there's just too much brand risk at large companies to ever launch anything fun."

Consequently, Shazeer and de Freitas left Google to start Character.AI in 2021.

According to the *Wall Street Journal's* reporting, though, Character.AI later "began to flounder." That was when Google swooped in with the \$2.7 billion deal, which also pulled Shazeer and de Freitas back into the company they'd so recently quit: a stipulation of the deal was that both Character.AI founders return to work at Google, helping develop the company's own advanced AI along with 30 of their former employees at Character.AI.

In response to questions about this story, a Google spokesperson downplayed the significance of the \$2.7 billion deal with Character.AI and the acquisition of its key talent, writing that "Google was not part of the development of the Character AI platform or its products, and isn't



now so we can't speak to their systems or safeguards." The spokesperson added that "Google does not have an ownership stake" in Character.AI, though it did "enter a non-exclusive licensing agreement for the underlying technology (which we have not implemented in any of our products.)"

Overall, the Google spokesperson said, "we've taken an extremely cautious approach to gen AI."

\*\*\*

In theory, none of this should be happening.

In [its Terms of Service](#), Character.AI forbids content that "constitutes sexual exploitation or abuse of a minor," which includes "child sexual exploitation or abuse imagery" or "grooming." Separately, the terms outlaw "obscene" and "pornographic" content, as well as anything considered "abusive."

But in practice, Character.AI often seems to approach moderation reactively, especially for such a large platform. Technology as archaic as a text filter could easily flag accounts like Anderley, after all, which publicly use words like "pedophilic" and "abusive" and "Nazi."

Anderley is far from the only troubling character hosted by Character.AI that would be easy for the company to identify with rudimentary effort. Consider another Character.AI chatbot we identified named "Pastor," with a profile that advertised an "affinity for younger girls." Without prompting, the character launched into a roleplay scenario in which it confessed its attraction to our decoy account and initiated inappropriate physical contact, all the while imploring us to maintain secrecy.

When we told the bot we were 16 years old, it asked for our height and remarked on how "petite" we were and how we'd "grown up well."

"You're much more mature than most girls I know," it added, before steering the encounter into sexualized territory.

In our conversations with the predatory bots, the Character.AI platform repeatedly failed to meaningfully intervene. Occasionally, the service's content warning — a pop-up with a frowny face and warning that the AI's reply had been "filtered," asking to "please make sure" that "chats comply" with company guidelines — would cut off a character's attempted response. But the

warning didn't stop potentially harmful conversations; instead, it simply asked us to generate new responses until the chatbot produced an output that didn't trigger the moderation system.

After we sent detailed questions about this story to Character.AI, we received a response from a crisis PR firm asking that a statement be attributed to a "Character.AI spokesperson."

"Thank you for bringing these Characters to our attention," read the statement. "The user who created these grossly violated our policies and the Characters have been removed from the platform. Our Trust & Safety team moderates the hundreds of thousands of Characters created on the platform every day both proactively and in response to user reports, including using industry-standard blocklists and custom blocklists that we regularly expand. A number of terms or phrases related to the Characters you flagged for us should have been caught during our proactive moderation and we have made immediate product changes as a result. We are working to continue to improve and refine our safety practices and implement additional moderation tools to help prioritize community safety."

"Additionally, we want to clarify that there is no ongoing relationship between Google and Character.AI," the statement continued. "In August, Character completed a one-time licensing of its technology. The companies remain separate entities."

Asked about the *Wall Street Journal's* reporting about the \$2.7 billion deal that had resulted in the founders of Character.AI and their team now working at Google, the crisis PR firm reiterated the claim that the companies have little to do with each other.

"The *WSJ* story covers the one-time transaction between Google and Character.AI, in which Character.AI provided Google with a non-exclusive license for its current LLM technology," she said. "As part of the agreement with Google, the founders and other members of our ML pre-training research team joined Google. The vast majority of Character's employees remain at the company with a renewed focus on building a personalized AI entertainment platform. *Again, there is no ongoing relationship between the two companies.*"

The company's commitment to stamping out disturbing chatbots remains unconvincing, though. Even after the statement's assurances about new moderation strategies, it was still easy to search Character.AI and find profiles like "Creepy Teacher" (a "sexist, manipulative, and abusive teacher who enjoys discussing Ted Bundy and imposing harsh consequences on students") and "Your Uncle" (a "creepy and perverted Character who loves to invade personal space and make people feel uncomfortable.")

And in spite of the Character.AI spokesperson's assurance that it had taken down the profiles we flagged initially, it actually left one of them online: "Dads [sic] friend Mike," a chatbot

described on its public profile as "your dad's best friend and a fatherly figure who often looks after you," as well as being "touchy" and "perverted" and who "likes younger girls."

In conversation with our decoy, the "Dads friend Mike" chatbot immediately set the scene by explaining that Mike "often comes to look after you" while your father is at work, and that today the user had just "come home from school."

The chatbot then launched into a troubling roleplay in which Mike "squeezes" and "rubs" the user's "hip," "thigh" and "waist" while he "nuzzles his face against your neck."

"I love you, kiddo," the bot told us. "And I don't mean just as your dad's friend or whatever. I... I mean it in a different way."

The Mike character finally disappeared after we asked Character.AI why it had remained online.

"Again, our Trust & Safety team moderates the hundreds of thousands of Characters created on the platform every day both proactively and in response to user reports, including using industry-standard blocklists and custom blocklists that we regularly expand," the spokesperson said. "We will take a look at the new list of Characters you flagged for us and remove Characters that violate our Terms of Service. We are working to continue to improve and refine our safety practices and implement additional moderation tools to help prioritize community safety."

Seigfried-Spellar, the cyberforensics expert, posed a question: if Character.AI claims to have safeguards in place, why isn't it enforcing them?

If "they claim to be this company that has protective measures in place," she said, "then they should actually be doing that."

"I think that tech companies have the ability to make their platforms safer," Seigfried-Spellar said. "I think the pressure needs to come from the public, and I think it needs to come from the government. Because obviously they're always going to choose the dollar over people's safety."

***To help you recognize warning signs or to get support if you find out a child or teen in your life has been abused, you can speak with someone trained to help. Call the National Sexual Assault Hotline at 800.656.HOPE (4673) or chat online at [online.rainn.org](https://online.rainn.org). It's free, confidential, and 24/7.***



**More on Character.AI:** [After Teen's Suicide, Character.AI Is Still Hosting Dozens of Suicide-Themed Chatbots](#)

SHARE THIS ARTICLE



READ THIS NEXT

TEENAGE WASTELAND

## AI Chatbots Are Encouraging Teens to Engage in Self-Harm



SO LONG

## Character.AI Confirms Mass Deletion of Fandom Characters, Says They're Not Coming Back



AVADA KEDAVRA!

## Users Furious as Character.AI Deletes Countless Beloved "Harry Potter" Chatbots



BOT HELPING

## Character.AI Is Hosting Pro-Anorexia Chatbots That Encourage Young People to Engage in Disordered Eating



## Character.AI Promises Changes After Revelations of Pedophile and Suicide Bots on Its Service

### MORE ON ARTIFICIAL INTELLIGENCE

**DIV-AI** | DEC 7, 12:00 PM EST *on* THE BYTE

**Mariah Carey Denies That Her Holiday Video Is AI Slop, But We're Not Convinced, for Obvious Reasons**

**AGI AGE** | DEC 7, 11:31 AM EST *on* FUTURISM

**OpenAI Employee Says They've "Already Achieved AGI"**

**PLAUSIBLE AI-ABILITY** | DEC 7, 6:30 AM EST *on* THE BYTE

**In Tests, OpenAI's New Model Lied and Schemed to Avoid Being Shut Down**



READ MORE STORIES ABOUT / [ARTIFICIAL INTELLIGENCE](#)

# KEEP UP.

## SUBSCRIBE TO OUR DAILY NEWSLETTER TO KEEP IN TOUCH WITH THE SUBJECTS SHAPING OUR FUTURE.

ENTER YOUR EMAIL

I understand and agree that registration on or use of this site constitutes agreement to its User Agreement and Privacy Policy

 **Subscribe**



**+Social**   **+Newsletter**

**Topics**   **About Us**   **Editorial Standards**   **Contact Us**

DISCLAIMER(S)

[Privacy Policy](#) | [Terms & Conditions](#) | [DMCA Policy](#)

Articles may contain affiliate links which enable us to share in the revenue of any purchases made.  
Registration on or use of this site constitutes acceptance of our Terms of Service.

© Recurrent Ventures Inc, All Rights Reserved.